



Using Historical Scientific Controversies to Promote Undergraduates' Argumentation

Pablo Antonio Archila¹  · Jorge Molina² · Anne-Marie Truscott de Mejía³

Published online: 27 May 2020
© Springer Nature B.V. 2020

Abstract

The Covid-19 pandemic is the reason why humanity is paying more attention to the importance of regular and rigorous handwashing. Interestingly, in the nineteenth century, regular and rigorous handwashing was a key (and controversial) solution proposed by the Hungarian obstetrician Ignaz Philipp Semmelweis to cut drastically cases of puerperal fever. The purpose of this study was to provide evidence that the case of Semmelweis and puerperal fever—a crucial historical scientific controversy—can be used as a springboard to promote university student argumentation. Our study was inspired by the fact that the Organization for Economic and Cooperative Development (OECD) stressed that more efforts and resources should be invested in promoting argumentation as an essential component for scientifically literate citizens in twenty-first century societies. However, nowadays, argument and debate are virtually absent from university science education. The data was derived from 124 undergraduates' (64 females and 60 males, 15–30 years old) written responses and audio and video recordings in a university biology course in Colombia. The findings show that the articulation of this historical controversy with decision-making, small-group debate, and whole-class debate activities can be useful for promoting undergraduates' argumentation. This study contributes to the development of a research-based university science education that can inform the design of an argumentation curriculum for higher education.

Keywords Argumentation · Decision-making · Historical scientific controversies · Semmelweis · Puerperal fever · University science education

✉ Pablo Antonio Archila
pa.archila@uniandes.edu.co

Jorge Molina
jmolina@uniandes.edu.co

Anne-Marie Truscott de Mejía
atruscot@uniandes.edu.co

Extended author information available on the last page of the article

1 Introduction

Argumentation is a crucial cognitive-linguistic skill required for the twenty-first century thinking and communicative citizen (Archila 2013; Archila 2015a; Archila et al. 2017). In parallel with this view, Andrews (2010, 2015) stresses the idea that argumentation in science is a chain of intellectual (and communicative) processes that should be practiced in the science classrooms at all educational levels—from primary to higher. In the recent *Dictionnaire de l'Argumentation* (Plantin 2016), a fundamental clarification of what argumentation implies is stated in the principle that arguing requires the construction of arguments in a (1) rational and (2) reasonable way. These two elements largely explain why the promotion of student argumentation is a complex and time-consuming process. Therefore, it should come as no surprise that the promotion of this skill requires practice. However, the problem is that argument and debate are virtually absent from university science education (Andrews 2010, 2015; Archila et al. 2018a, 2018b, 2018c, 2019; Pabuccu and Erduran 2017; Wieman 2017). The reason is straightforward: the main teaching and learning actions in most university science courses around the world focus solely on undergraduates' acquisition of scientific conceptual content (Archila et al. 2018a, b, c).

Clearly, if a student is never challenged to make informed decisions, argue, and debate, he/she will never cultivate the skills to participate appropriately in the communicative practices of science. Accordingly, the Organization for Economic and Cooperative Development (OECD) (OECD 2017) has stressed that more efforts and resources should be invested in promoting argumentation as an essential component for scientifically literate citizens in twenty-first century societies. In this article, “promoting students' argumentation is understood as the opportunity for learners to build arguments related to a decision (in our case, about a historical scientific controversy) made by themselves” (Archila 2015b, p. 1203).

The use of historical scientific controversies to promote undergraduate argumentation is an under-researched possibility in higher science education (Adúriz-Bravo 2014; de Hosson 2011; Garritz 2013; Justi and Mendonça 2016; Zemplén 2011). Indeed, there is not much evidence of their use to promote argumentation in higher science education (Archila 2015b). In response to the need to promote argumentation in university science courses, in this study, we propose a teaching-learning sequence (TLS) based on the case of Semmelweis and puerperal fever—a crucial historical scientific controversy. This empirical study focuses on providing evidence to show that the case of Semmelweis and puerperal fever can be used as a springboard to promote university students' argumentation. In light of the goal of the current study, the research questions that guided this investigation were as follows:

- (1) Does the case of Semmelweis and puerperal fever cause controversy in a group of university students?
- (2) Does the TLS engage a group of university students in argumentative classroom interactions?

2 Theoretical Framework

In this section, we present the conceptual bases of the TLS proposed in this study. We used the term “teaching-learning sequence” (TLS) to refer to the articulation between proposed teaching and expected student learning as a distinguishing feature of such research-inspired subject-

oriented sequences (Psillos and Kariotoglou 2016). It is important to bear in mind that “a TLS is both an interventional research activity and a product” (Psillos 2015, p. 1036). Additionally, Psillos (2015) reminds us that a TLS can be a one-session class or alternatively can last various weeks. The TLS proposed in this study was designed as a single 80-min class session, dealing with the use of historical scientific controversies to promote undergraduate argumentation. Hence, the conceptual bases of this pedagogical strategy focused on six elements, namely: (1) argumentation, (2) decision-making, (3) historical scientific controversy, (4) argumentative interaction, (5) small-group debate, and (6) whole-class debate.

It is important to start by clarifying that in the context of the present study, the expression “claim” is used to indicate the students’ position (posture) about the historical controversy. The term “evidence” refers to the data and facts students use in support of their position. The word “argument” is understood as the product of the articulation (in a rational and reasonable way) of this evidence with the claim. Several authors (e.g., Andrews 2010; Archila 2018) maintain that decision-making is an effective means of creating student motivation and interest in argumentative practices. One reason for this is that it helps students feel their points of view have an important place in their educational process.

By a historical scientific controversy, we mean a historical scientific issue that will lead to a high level of different understandings among significant numbers of people (de Hosson and Kaminski 2007). According to de Hosson and Kaminski (2007) and de Hosson (2011), the use of historical scientific controversies in the science classroom can be understood as a means for science educators to provide students with opportunities to formulate arguments and to make decisions. In line with this possibility, Archila (2015b) recommends that the historical scientific controversy should be presented to the students in the form of a controversial question. This is through a (1) provoking and (2) ambiguous question. The reason for this is that these two characteristics are relevant to help students focus on the pieces of evidence they use rather than the decisions they make about the controversial question.

Furthermore, in order to promote argumentation, it is imperative to increase student voice in the science classroom. Archila et al. (2018a), Muller Mirza (2015), and Schwarz and Baker (2017) pointed out the fact that student voice is increased when they are engaged in argumentative interaction (such as small-group debate and whole-class debate). Moreover, Baker (2002, 2009) stressed the idea that controversy is the bedrock of argumentative interaction. Thus, we define argumentative interaction as a dialogue—around a controversial question—between two or more individuals in which they communicate a claim, provide evidence, and articulate this evidence with the claim to produce arguments and thus make an informed decision. In this study, we adopted the five conditions for argumentative interaction suggested by Baker (2002, 2009). These can be briefly summarized as follows: (1) the existence of a diversity of proposals relating to a controversy, (2) different proposals should exist in the same small group, (3) each proposal should be acceptable (reasonable), (4) each small group is asked to choose one proposal, and (5) when making a decision, each small group should carefully examine the arguments of the selected proposal.

3 Literature Review

Having detailed the conceptual bases of our TLS, in this section, we will now discuss previous studies interested in exploiting the possibility of connecting historical controversies and argumentation. Additionally, we explain why the case of Semmelweis and puerperal fever is

an under-researched option to promote university students' argumentation. The literature review in this section is important to support the claim that the significance of the present study lies in the fact that it expands on the scope of some notable work carried out previously that has focused on the promotion of student argumentation.

3.1 Connecting Historical Controversies and Argumentation

Much of the efforts to include historical controversies in science classrooms have focused on providing more informed views of the nature of science (Nouri and McComas 2019). Indeed, there is little evidence of the use of this type of controversy to promote argumentation in university science courses (Archila 2015b). Recently, Justi and Mendonça (2016) used the controversy about the awarding of the Nobel Prize in Chemistry to Fritz Haber in 1918. As part of a teacher training project in Brazil, 16 future chemistry teachers participated in a dramatization activity, in which they discussed this controversy. According to Justi and Mendonça (2016), this activity was not only meaningful in promoting participants' argumentation and informed views of nature of science but also in fostering reflection on their future actions related to the authentic teaching of and about science.

In France, Archila (2015c) presented a group of 63 high school students with the controversial question of who discovered oxygen, with Carl Wilhelm Scheele (1742–1786), Joseph Priestley (1733–1804), and Antoine Laurent de Lavoisier (1743–1794) being possible candidates. Participants were asked to evaluate evidence relating to this controversy. He found that the historical controversy “Who discovered oxygen?” appears to be particularly promising in terms of encouraging students to evaluate evidence relating to experimentation in science and scientific communication. In Brazil, de Oliveira and Mendonça (2019) used the case of the historical controversy of oxygen gas as a crucial aspect of a pedagogical proposal to explicitly promote the argumentation of seven pre-service chemistry teachers. In a similar vein, Zemplén (2011) claims that there is much work to be done in this area and further research should explore the vast multiplicity of historical scientific controversies in order to promote student argumentation in different parts of the world.

Despite the fact that in higher education little is known about the use of an approach combining historical scientific controversy and argumentation (Archila 2014; Garritz 2013), it is rational and reasonable to explore the possibility that in a university science course, the instructor should consider the use of a historical controversy to involve undergraduates in decision-making and thus to take advantage of this activity to encourage and facilitate argumentative interaction (such as small-group debate and whole-class debate). In this way, students will be exposed to the decisions of others, as well as sources of evidence, viewpoints, and reasoning. In the next section, we explain why the case of Semmelweis and puerperal fever is an under-researched option to promote university student argumentation.

3.2 The Case of Semmelweis and Puerperal Fever

Puerperal (or childbed) fever is an infection of some part of the female reproductive organs following childbirth or abortion. This infection often occurs during the puerperium (approximately 6 weeks after childbirth) when the womb returns to its normal shape. Fever of 311.15 K (38 °C) and higher during the first 10 days following delivery or miscarriage is a key symptom. Puerperal infection is most commonly found in the raw surface of the interior of the uterus after separation of the placenta (afterbirth), but pathogenic organisms may also

affect lacerations of any part of the genital tract. They can invade the bloodstream and lymph system to cause cellulitis (inflammation of the cellular tissue), peritonitis (inflammation of the abdominal lining), and septicemia (blood poisoning) (Encyclopædia Britannica 2018).

Loudon (1986) reminds us that the first recorded epidemic of puerperal fever occurred at the Hôtel Dieu in Paris in 1646. Afterward, maternity hospitals all over Europe and North America reported intermittent outbreaks, and even between epidemics, the death rate from sepsis reached one woman in four or five of those giving birth. Puerperal fever was often fatal, with death usually occurring 5 to 10 days after delivery (Aragón-Méndez et al. 2018; Ataman et al. 2013; Best and Neuhauser 2004). The year 2018 marked the 200th anniversary of the birth of the Hungarian obstetrician Ignaz Philipp Semmelweis (or Ignác Fülöp Semmelweis, also known as the “savior of mothers” and “father of infection control”) (1 July 1818–13 August 1865). He was the first to identify the mode of transmission of puerperal fever. According to Kadar (2019), no other obstetrician has had so many honors showered on him after his death or been treated so unjustly during his lifetime as Semmelweis. His history and investigations into puerperal sepsis are some of the most interesting in the history of science. These have been detailed by Carter and Carter (2017), Nuland (2003), Obenchain (2016), and others.

For various reasons, Semmelweis’s ideas failed to be considered as rational and reasonable until after his death. Rather than evaluating with fair-mindedness Semmelweis’s strong evidence of a clear lifesaving intervention, obstetricians in Vienna and elsewhere disputed it (Lerner 2014; Stewardson and Pittet 2011). The first possible reason for Semmelweis’s failure has to do with his nationality. It should be noted that 1848–1850 was a time of great social unrest within the Austro-Hungarian Empire. Consequently, as a Hungarian belonging to a German-speaking ethnic minority in Vienna, Semmelweis was an ardent believer in wishing to see Hungary emerge from its second-class status. In this respect, some authors (e.g., Dunn 2005; Gillies 2005; Henao 1999; Stewardson and Pittet 2011; Volcy 2012) believe that critics did not evaluate Semmelweis’s evidence fairly. In other words, they were biased because of his nationality.

With respect to the second tentative reason, Loudon (2013) highlights the fact that Semmelweis waited for 13 years before he published (in 1860) his treatise, *The Etiology, Concept, and Prophylaxis of Childbed Fever*. The situation becomes even more complicated because the treatise of over 500 pages contains passages of great clarity interspersed with lengthy, muddled, repetitive, and bellicose passages in which he attacks his critics. In line with this focus on the relevance of communication to scientific progress, Dunn (2005), Gillies (2005), Lerner (2014), Stewardson and Pittet (2011), and Volcy (2012), among others, agree that in part, Semmelweis’s failure to achieve recognition for his contribution was due to his own failure to publicize his findings.

The third plausible reason emerges from Semmelweis’s advice for medical caregivers about washing their hands in calcium chloride before examining women who had recently given birth. Lerner (2014, p. 210) claims that many Semmelweis’s colleagues considered as “heretical” the notion that puerperal fever was spread by medical personnel through direct physical contact. This advice was regarded by Semmelweis’s fellow physicians as implying that they were dirty. Far from being grateful for the huge reduction in fatal puerperal fever cases, they labeled Semmelweis a crackpot who was insulting their honor (Cropley and Cropley 2008). Accordingly, this would explain why Semmelweis’s advice met a lot of resistance from his colleagues (Adriaanse et al. 2000). Also, Stewardson and Pittet (2011) and Volcy (2012) stress that washing medical personnel’s hands the right way in calcium chloride was a time-

consuming activity (an interminable 5 min) which caused much irritation to the hands. For this reason, they consider that this was unpopular advice for such busy people as Semmelweis's fellow physicians. Volcy (2012) asserts that a pedagogic campaign led by Semmelweis would have helped attendants set aside their dominant position in obstetrics and become aware of the importance of disinfecting their hands with a calcium chlorine solution before examining women in labor.

The fourth and last possible reason for Semmelweis's failure concerns the fact that Semmelweis's views clashed with the dominant paradigm of his day: the miasma and contagion theories (Gillies 2005; Hernández Botero 2010; Persson 2009; Volcy 2012). Although eminently rational and reasonable, Semmelweis' discovery directly confronted the beliefs of science and medicine in his time (Ataman et al. 2013; Hernández Botero and Florián Pérez 2012; Kadar 2019). Henao (1999) describes Semmelweis' contribution as the genesis of a new paradigm: the germ theory of disease. This new paradigm was resisted because it flew in the face of all existent theories held by prominent obstetricians of his day (Kadar 2019). Moreover, as Gillies (2005) reminds us, the germ theory of disease only began to take shape following the researches of Pasteur in the late 1850s and 1860s. It only became generally accepted in the medical community in the 1880s. All this was in time for Lister but too late for Semmelweis.

3.3 The Teaching-Learning Sequence

As mentioned in the conceptual framework section, the conceptual bases of our TLS focused on the following six elements: (1) argumentation, (2) decision-making, (3) historical scientific controversy, (4) argumentative interaction, (5) small-group debate, and (6) whole-class debate. Moreover, the TLS consists of four steps (Table 1) and has four main characteristics: (1) it provides university students with opportunities to make decisions about a controversial question; (2) at each step, undergraduates have the opportunity to enrich the positions they use to argue in favor of their decision; (3) they can change the decision if and whenever they want to; and most importantly, (4) each step has been designed to consistently engage students in argumentation (Archila 2015b).

Archila (2015b) considers that formulating a controversial question is a key component of a TLS that aims to promote students' argumentation. In this TLS, the controversial question, "Why did Semmelweis's views not gain acceptance among many members of the medical community in his time?" was inspired by the historical controversy described in the narrative, Semmelweis and puerperal fever, elaborated by Acevedo-Díaz et al. (2016). Possible answers were as follows: (a) Semmelweis's nationality, (b) Semmelweis's failure to publicize his findings, (c) physicians' reluctance to washing their hands in calcium chloride, and (d) Semmelweis's views clashed with the dominant paradigm.

"Why did Semmelweis's views not gain acceptance among many members of the medical community in his time?" This is a controversial question, for two reasons: (1) this question may cause classroom discussion, and (2) there could be strong (rational and reasonable) student arguments for each possible answer (Archila 2015b). It is important to clarify that in this TLS, the pieces of evidence ("Why did you make that decision?") that undergraduates may use are more important than the decisions they make about the question. In other words, there is no one right answer to the controversial question (Acevedo-Díaz et al. 2016). The TLS was designed as a single 80-min class session. All four steps of the TLS were organized around the controversial question (Table 1). The TLS begins by inviting students to answer

this question (step 1 in Table 1). Students developed informed opinions individually (steps 1 and 4) and in groups (steps 2 and 3) with a view to answering the controversial question.

3.4 The Narrative Presented to the Undergraduates

Semmelweis and puerperal fever is a narrative of 2450 words, created and proposed by Acevedo-Díaz et al. (2016) as a didactic resource to foster informed views of the nature of science. The promotion of students' understanding of the nature of science is not the focus of our study. That said, in the present study, this narrative has been used for the first time to explicitly promote students' argumentation. Indeed, we decided to present this narrative to the university students for five reasons: (1) during the process of development, Acevedo-Díaz et al. (2016) used several literature sources to produce as holistic a narrative as possible; (2) the narrative revolves around the controversy, "Why did Semmelweis's views not gain acceptance among many members of the medical community in his time?"; (3) in the narrative, plausible reasons for Semmelweis's failure are explained; (4) Aragón-Méndez et al. (2016, 2018) stress that satisfactory results have been obtained when this narrative has been implemented in educational activities with secondary students and prospective biology teachers; and (5) Acevedo-Díaz et al. (2016) developed the narrative in a way that requires no special background in biology, in medicine, or in the history of science.

3.5 The Role of the Instructor

In this TLS, as recommended by Archila (2015b), we acknowledge that the promotion of university students' argumentation implies moving away from the instructor's role as the unquestioned authority providing all the answers in the class. Therefore, the instructor implemented the strategy, acting as a facilitator in steps 1 and 4, and perhaps more importantly, that of a challenger in steps 2 and 3 (Table 1). He made evaluative comments in response to the students' argumentative interaction to help them criticize the argumentation of others. While doing this, the instructor took care not to influence students' decisions by demonstrating and maintaining his neutrality throughout—whatever values the undergraduates defended had to be accepted and respected. In other words, throughout the four steps of the TLS, his sole function was to encourage the university students and engage them discussing and evaluating the tentative reasons for Semmelweis's failure.

4 Research Design and Method

4.1 Context and Participants

The TLS was implemented in a university bilingual (Spanish-English) science course (Archila et al. 2018a) called Biology of Organisms. This course was chosen by convenience sampling (Bryman 2016). Much of the reason for this is that the second author is the course instructor. Biology of Organisms is a large (75–95 students per semester), introductory course that is offered every semester by the Department of Biological Sciences to participants in all undergraduate programs at a private university in Bogotá, Colombia. This university has a high academic ranking in Latin America. Its educational policy is to foster the integration of students from different majors and different age groups. Thus, it is very common to see

Table 1 Teaching-learning sequence

| Step | Activity | Details | Didactic Purpose |
|------|---|--|--|
| 1 | Questionnaire, Part One: Making an initial decision | <p>The undergraduates, individually, read the narrative (2450 words), Semmelweis and puerperal fever (Acevedo-Díaz et al. 2016). Then, three students read aloud assuming the roles of storyteller, Semmelweis, and footnote reader</p> <p>Each student uses the narrative provided to identify the claim, evidence, and arguments used by Semmelweis, thereby enabling each student to argue for or against proposals and arrive at a decision on an individual basis</p> <p>Each student answers the question, “Why did Semmelweis’s views not gain acceptance among many members of the medical community in his time?”</p> | <p>The didactic purpose of Part One of the questionnaire is to encourage each undergraduate to explore Semmelweis’ argumentation (claim, evidence, and arguments) and make an initial decision (before debating)</p> <p>The instructor introduces the class to the context of the controversial question, “Why did Semmelweis’s views not gain acceptance among many members of the medical community in his time?”</p> <p>The controversial question posed to each student has four possible answers: (a) Semmelweis’s nationality, (b) Semmelweis’s failure to publicize his findings, (c) physicians’ reluctance to washing their hands in calcium chloride, and (d) Semmelweis’s views clashed with the dominant paradigm. Why did you make this decision?</p> |
| 2 | Argumentative interaction (small-group debate) | <p>The undergraduates are organized in groups of three or four to discuss the decision made by each undergraduate in step 1 with a view to making a group decision</p> | <p>Step 2 has been designed to explicitly promote argumentative interaction. The discussion among undergraduates enables them to learn about the decision made by each group member in step 1. Students’ understanding of Semmelweis’s argumentation is crucial to the discussion</p> <p>Making a group decision constitutes a challenge that will motivate students’ argumentation during this step which involves deliberation</p> |
| 3 | Argumentative interaction (whole-class debate) | <p>The entire class debates, taking into account the decisions made by each group in step 2</p> <p>Each group announces its decision to the class and presents the underlying arguments</p> | <p>The different reasons described in the narrative constitute a transcendent point of the whole-class debate, as in steps 1 and 2</p> <p>The instructor’s intervention is highly relevant in promoting argumentation in the debate: as each group communicates its decision, the teacher asks open-ended questions, based on the arguments produced by the students, to challenge these arguments</p> |
| 4 | Questionnaire, Part Two: Making a final decision | <p>To conclude the debate held during step 3, each undergraduate writes down his or her final decision about the question, “Why did Semmelweis’s views not gain acceptance among many members of the medical community in his time?”</p> | <p>The didactic purpose of step 4 is to provide the opportunity for a final reflection (based on the preceding steps) to encourage the making of a final decision</p> <p>Before this step, each undergraduate would have communicated and argued his or her position (once in writing during step 1 and once orally in step 2) on the controversial question. In addition, each small group would have expressed and argued its group decision (step 3)</p> |

students from different socioeconomic status, academic achievement, majors (not only Biology and Microbiology), and ages taking this bilingual course. Syllabus contents include the following: Biomolecules, Evolution, Phylogenetics, Systematics, Archaea, Bacteria, Eukaryotes, Seedless plants, Seed plants, Fungi, Protostome, Ecdysozoa, and Ambulacraria.

Among the 142 eligible students, 124 (87.3%) participated in this study. Out of these 124 participants, 64 (51.6%) were females, and 60 (48.3%) were males. The age distribution ranged from 15 to 30 years, and the average age was 19.1 years ($SD = 2.05$). The authors informed the undergraduates that their answers would have no influence on their final course grade and that they could withdraw at any time. The reason for this is that, as pointed out by Bryman (2016), this is a way to facilitate participants' spontaneity and naturalness and reduce anxiety. He also stresses that there is a loss of spontaneity when students feel they are being assessed.

Participants and their parents were informed of the general research purpose. All responses were kept confidential. The authors ensured that the inquiry was not harmful to any participant involved. Specifically, harm to participants' emotional, intellectual, and social development and loss of self-esteem was avoided through the generation of a climate of confidence and respect in the science classroom (Bryman 2016). All participants were treated in accordance with the ethical guidelines of the American Psychological Association (APA) with respect to consent, confidentiality, and anonymity. For this reason, the undergraduates were assigned codes to protect their privacy, for example, 1U46 means Class 1, undergraduate number 46.

These 124 university students were grouped into two classes. The TLS was carried out in the following order:

Class 1: Undergraduates taking Biology of Organisms during the first semester (average age 19.1 years), 69 students (35 females and 34 males).

Class 2: Undergraduates taking Biology of Organisms during the second semester (average age 20.8 years), 55 students (29 females and 26 males).

We decided to implement the TLS in two classes to increase our sample size. It is important to clarify that data from the two classes was analyzed independently. Much of the reason for this is that the sequence was implemented during the first (Class 1) and second (Class 2) semester of the same academic year.

4.2 Data Collection

Data were collected from written responses and audio and video recordings. The written responses were obtained by means of a pen and paper questionnaire (Appendix 1) completed by all 124 participants in steps 1 and 4 (Table 1), while audio and video recordings were obtained from the undergraduates' small-group debates in step 2 and the whole-class debate in step 3 (Table 1). The questionnaire and the narrative, Semmelweis and puerperal fever (Acevedo-Díaz et al. 2016), were distributed to the participants at the beginning of the TLS. The whole questionnaire and the narrative were printed in Spanish. Participants were given the option to decide the language (Spanish, English, or a hybrid version using code-switching) they wanted to use for writing their answers to each question due to the Spanish-English bilingual nature of the university science course (Archila et al. 2018a). This questionnaire contained five questions presented in two parts (Appendix 1).

At home, participants were given 5 days to read the narrative and answer the first three questions. The aim of these three questions was to facilitate the undergraduate students' reading comprehension (the results relating to these questions are not discussed in this article).

Then, in the classroom, three students read aloud the narrative assuming the roles of storyteller, Semmelweis, and footnote reader. The students answered the same questionnaire (“Why did Semmelweis’s views not gain acceptance among many members of the medical community in his time?”) in steps 1 (Question 4) and 4 (Question 5) of the TLS (Table 1). Students were given 15–20 min to reread the narrative and answer Question 4 independently (Appendix 1). Each student had a copy of the text of the scene to which he/she could refer during the TLS. In the second part of the questionnaire, the students were given 5–10 min to make a final decision about the controversial question (step 4 in Table 1).

Four stereo digital voice recorders were set up for Classes 1 (69 participants) and 2 (55 participants) in order to record the students’ small-group debates in step 2 (25–30 min) and the whole-class debate in step 3 (15–20 min). Additionally, one video camera was placed in each of the two classrooms. Each small-group debate was conducted among three or four students who were used to working together during class activities. Archila (2015b, 2017) maintains that a TLS should be understood as a continuous refinement process. To be precise, he recommends generating a climate of confidence in the science classroom and asking students for key points that help the researchers continuously improve the sequence. This is why to find out participants’ opinions about the TLS, at the end of the implementation, they were asked to answer a pen and paper survey (5–10 min) (Appendix 2). It is important to clarify that even although analysis of the students’ opinions was not explicitly included in any of our two research questions, we decided to assume this opinion as a valuable feedback for future improvements of our student-centered sequence. We created the survey based on questions previously formulated by Archila (2015b, 2017). For this reason, we considered that these were valid to find out about participants’ opinion of the TLS. Participation in the survey was completely anonymous. The whole survey was printed in Spanish, and participants were given the option to decide which language (Spanish, English, or a hybrid version using code-switching) they wanted to use for writing their answers to the open-ended questions.

4.3 Data Analysis

Data analysis was carried out at two levels in order to answer the two research questions: (1) analysis of participants’ decisions about the historical controversy and (2) analysis of participants’ argumentative interaction. The first research question, “Does the case of Semmelweis and puerperal fever cause controversy in a group of university students?” was addressed quantitatively using frequency counts. In other words, the focus of this first level of analysis was the students’ decisions related to the question, “Why did Semmelweis’s views not gain acceptance among many members of the medical community in his time?”, in steps 1 (initial decision) and 4 (final decision) of the TLS. The calculation of the frequency counts of the four plausible reasons (Semmelweis’s nationality, Semmelweis’s failure to publicize his findings, physicians’ reluctance to wash their hands in calcium chloride, and Semmelweis’s views clashing with the dominant paradigm) helped us to find out whether (or not) the case of Semmelweis and puerperal fever caused controversy among the participants. As part of the continuous refinement process and to have an idea of the participants’ opinion of our TLS, their responses to the survey (Appendix 2) were analyzed using frequency counts (Question 1, 2, 4, 7, 8, 9 in Appendix 2). Some answers to open-ended questions (Questions 3, 5, 6 in Appendix 2) are commented on in the Results section.

Our second research question, “Does the TLS engage a group of university students in argumentative classroom interactions?” was addressed using the qualitative data analysis

software Transana® (Mavrou et al. 2007), which was used to code the transcripts of the video and audio recordings of the episodes in which the small-group debate and the whole-class debate (steps 2 and 3 in Table 1) took place. Transana® (transana.com) is a program conceptualized and developed by the Wisconsin Centre for Educational Research at the University of Madison. This software was designed to handle large audio and video collections and to facilitate the documentation of analysis and the interpretation of data. Within the multiple tools of this program, three were especially beneficial for our study, namely: (1) simple keyboard shortcuts for controlling our media files from within the transcription window, resulting in a process of manual transcription as fast and efficient as possible; (2) combination of both audio and video with text analysis (transcripts); and (3) transcripts can be time coded and synchronized with the same piece of audio or video, giving us the possibility to create coding schemes for different observations (in our case argumentative oral interaction) which correspond to extracts from the transcripts.

In accordance with verbal protocol analysis—a data collection and analysis method used to make valid inferences from transcripts (Ruiz-Primo 2015)—the episodes were transcribed verbatim using Transana® and then proofread to enhance the rigor and quality of the transcripts. After that, we created the following four coding schemes: (1) Semmelweis's nationality, (2) Semmelweis's failure to publicize his findings, (3) physicians' reluctance to wash their hands in calcium chloride, and (4) Semmelweis's views clashing with the dominant paradigm. These four schemes were created according to the evidence communicated in the narrative, Semmelweis and puerperal fever (Acevedo-Díaz et al. 2016), that the students used to argue in favor of their decisions. Some transcripts are commented on in the Results section. Strictly speaking, the role of these four schemes was to facilitate the documentation of the students' engagement in argumentative classroom interactions (small-group debate and whole-class debate). This documentation helped us to make better interpretations and valid inferences of data from the transcripts. These transcripts are English translations of the verbatim Spanish transcripts, care having been taken to remain as faithful as possible to the original meanings and wording.

5 Results

The results of the implementation of the TLS are presented in the following two sections; the first section presents the results of the questionnaire (steps 1 and 4 in Table 1) in response to our first research question, while the second section presents the results of students' engagement in argumentative interaction (steps 2 and 3 in Table 1) in response to our second research question. The results of the survey are presented in both two sections to provide a deeper context for the results of each of the four steps.

5.1 Results of Undergraduates' Responses to the Questionnaire

de Hosson and Kaminski (2007) remind us that a historical scientific controversy is a historical scientific issue that leads to a high level of different understandings among significant numbers of people. With this in mind, in our TLS, the initial decision (step 1) and the final decision (step 4) are related to our first research question, "Does the case of Semmelweis and puerperal fever cause controversy in a group of university students?". Thus, in steps 1 and 4, each student answered the question, "Why did Semmelweis's views not gain acceptance among

many members of the medical community in his time?”. Table 2 shows the four plausible reasons for Semmelweis’s failure (options A, B, C, and D) and the decisions made during steps 1 and 4 and by each class. Specifically, this table shows that in step 1 (initial decision), in both classes, a high number of participants (52/69 in Class 1; 42/55 in Class 2) considered the fact that Semmelweis’s failure to publicize his findings (option B) affected negatively the acceptance of his viewpoints among many of the members of the medical community in his time. Options A (35/69 in Class 1; 32/55 in Class 2) and D (33/69 in Class 1; 30/55 in Class 2) caused a clear discrepancy among participants.

It is interesting to note that very few students (4/69 in Class 1; 5/55 in Class 2) considered that physicians’ reluctance to washing their hands in calcium chloride (option C) could be a reason to explain why Semmelweis’s viewpoints did not gain acceptance among many of the members of the medical community in his time. This is interesting for two reasons. First, some participants were implicitly assuming the role of these physicians of the nineteenth century who openly showed their reluctance to washing their hands in calcium chloride. Second, some students did not attach much value to the possibility that many of Semmelweis’s colleagues considered “heretical” (Lerner 2014, p. 210) the notion that puerperal fever was spread by medical personnel through direct physical contact. Hence, this controversy among students is a contribution to the answer to our first research question.

In step 4 (final decision), Table 2 indicates that a representative number of participants considered that options A (49/69 in Class 1; 44/55 in Class 2), B (67/69 in Class 1; 52/55 in Class 2), and D (52/69 in Class 1; 39/55 in Class 2) answered the controversial question, “Why did Semmelweis’s views not gain acceptance among many members of the medical community in his time?”. In other words, the TLS helped participants to reflect that the articulation of more than one plausible reason (option) could explain Semmelweis’s failure. Also, Table 2 indicates that in both classes, there was an evident increase in the number of students who considered options A (from 35 to 49 in Class 1; from 32 to 44 in Class 2), B (from 52 to 67 in Class 1; from 42 to 52 in Class 2), and D (from 33 to 52 in Class 1; from 30 to 39 in Class 2) as possible reasons for Semmelweis’s failure. These results suggest that the level of discrepancy among participants identified in step 1 was reduced in step 4. Arguably, there is value to be gained in terms of argumentative interaction if discrepancy is treated as a resource in the classroom. One reason for this is that discrepancy can be used as a springboard for promoting

Table 2 Decisions made during steps 1 and 4 and by each class

| | Step 1 ($n = 124$) | Step 4 ($n = 124$) |
|----------------------|----------------------|----------------------|
| Class 1 ($n = 69$) | | |
| Option A | 35 | 49 |
| Option B | 52 | 67 |
| Option C | 4 | 5 |
| Option D | 33 | 52 |
| Class 2 ($n = 55$) | | |
| Option A | 32 | 44 |
| Option B | 42 | 52 |
| Option C | 5 | 7 |
| Option D | 30 | 39 |

Option A Semmelweis’s nationality, *Option B* Semmelweis’s failure to publicize his findings, *Option C* physicians’ reluctance to wash their hands in calcium chloride, *Option D* Semmelweis’s views clashed with the dominant paradigm

argumentative interaction through the small-group debate and the whole-class debate (steps 2 and 3).

As part of step 1, the students read the narrative, Semmelweis and puerperal fever (Acevedo-Díaz et al. 2016). This narrative offered participants the opportunity to familiarize themselves with four plausible reasons (options A, B, C, and D) to make an initial decision about the controversial question. This is why in the first step of our TLS, the narrative, Semmelweis and puerperal fever (Acevedo-Díaz et al. 2016), played a crucial role. To provide a deeper context, bearing in mind the importance of this narrative in our intervention, it is relevant to mention that the results of the pen and paper survey (Appendix 2) reveal that a significant number of respondents (68/69 in Class 1; 54/55 in Class 2) considered that the narrative, Semmelweis and puerperal fever (Acevedo-Díaz et al. (2016), was easily understandable for them (Question 3 in Appendix 2). Some comments include the following reasoning: “It explained both the social and the scientific context. Semmelweis’s steps during his research are described, so this was a clear explanation”, “the text was well-written, its lexicon was easy to understand, and the story of Semmelweis was narrated in a simple and concise way”, “even though it has medical vocabulary it is not complex”, and “the text has a logical order, its ideas are easy for the reader to understand, and the foot notes are useful to understand the context of the story a little more thoroughly”.

These comments can be explained by the fact that, as Acevedo-Díaz et al. (2016) assert, the narrative was constructed in a way that requires no special background in biology, in medicine, or in the history of science. Moreover, all the participants (69 in Class 1 and 55 in Class 2) acknowledged that they had had sufficient time for reading (Question 4 in Appendix 2). It is important to remember that in the present study, this narrative has been used for the first time to explicitly promote undergraduate argumentation. Furthermore, to better interpret the outcomes of the steps 1 and 4, the survey shows that very few participants (4/69 in Class 1; 0/55 in Class 2), apart from the Biology of Organisms course, had ever heard about Semmelweis’s contributions (Question 1 in Appendix 2). In contrast, a high number of students (63/69 in Class 1; 49/55 in Class 2) had received instruction in argumentation (Question 2 in Appendix 2).

5.2 Results Relating to Undergraduates Engagement in Argumentative Interaction

Steps 2 (small-group debate) and 3 (whole-class debate) are related to our second research question, “Does the TLS engage a group of university students in argumentative classroom interactions?”. In these steps, the controversial question, “Why did Semmelweis’s views not gain acceptance among many members of the medical community in his time?” was used as a platform to foster argumentative interaction. Table 3 shows the four plausible reasons for Semmelweis’s failure (options A, B, C, and D) and the decisions made by the undergraduates during the small-group debate (step 2). As mentioned previously, each group was composed of three or four students. Hence, there were nineteen small groups in Class 1 and fifteen in Class 2. During step 1 (initial decision), participants made a decision individually, and then in step 2, they interacted in small groups to make a group decision. The communication of each small-group debate decision marked the beginning of the whole-class debate (step 3). According to the results in Table 3, in both classes, there was evidence of a similar trend to the one reported in Table 2 (steps 1 and 4). This shows that a considerable number of small groups (19 in Class 1 and 13 in Class 2) considered the fact that Semmelweis’s failure to publicize his findings (option B) negatively affected the acceptance of his views among many of the members of the

Table 3 Decisions made during the small-group debate

| Decision | Class 1 (<i>n</i> = 19) | Class 2 (<i>n</i> = 15) |
|----------|--------------------------|--------------------------|
| Option A | 13 | 11 |
| Option B | 19 | 13 |
| Option C | 2 | 1 |
| Option D | 14 | 10 |

Option A Semmelweis's nationality, *Option B* Semmelweis's failure to publicize his findings, *Option C* physicians' reluctance to wash their hands in calcium chloride, *Option D* Semmelweis's views clashed with the dominant paradigm

medical community in his time, while options A (13/19 in Class 1 and 11/15 in Class 2) and D (14/19 in Class 1 and 10/15 in Class 2) caused more discrepancy among the small groups.

Step 2—Students' Engagement in Argumentative Interaction through Small-Group Debate

To reiterate, step 2 had been designed to explicitly promote argumentative interaction. During this step, the students showed signs of argumentative interaction when expressing additional information and reasoning relating to their decisions (Baker 2002, 2009). To illustrate this assertion, consider the following example:

1U1: Well, I decided "B" because I consider Semmelweis's failure was caused by the fact that he didn't publicize his results properly.

1U62: [interrupts] there is a part in the text which says that his [referring to Semmelweis] publication was a lengthy, repetitive, and sometimes confusing text as well as being difficult to read. So, it was difficult to understand his ideas [...].

1U29: I decided "A" and "D" taking into account the political context of that time; many physicians didn't support him [referring to Semmelweis] because they thought that behind his intentions there were certain ideologies. And then the [option] "D" because at that time dominant theories were unquestioned, as then the fact of questioning them was like creating a revolution in the scientific field [...].

1U55: I decided "D" because his [referring to Semmelweis] ideas always were clashing with what was thought in his time.

1U29: [interrupts] Yes, exactly. Also, some physicians reacted negatively to the idea of washing their hands because of laziness. Moreover, they were not aware of its importance.

1U62: However, I think that more than laziness it was like arrogance. Right?

1U29: Yes.

1U62: [interrupts] I mean, it was more like they were trying to say no, no, no, we are not going to wash our hand because we know that we aren't the ones who are spreading [the illness].

1U29: Exactly, because his [referring to Semmelweis] explanation went against dominant explanations at that time [...].

This excerpt was obtained from Transana® that allowed us not only to find out about the decision of each member of this small group but also to get access to how students interacted argumentatively when discussing the decision made by each undergraduate in step 1 with a view to making a group decision. Additionally, this excerpt indicates that the discussion among these undergraduates enabled them to learn about the decision made by the group members in step 1. 1U62's decision is not communicated explicitly as those of 1U1 (option B), 1U29 (options A and D), and 1U55 (option D). Arguably, having a small group in which its members made different decisions individually in step 1 can be assumed, according to Archila

(2015b, 2017) and Archila et al. (2019), as a good predictor of how successful that small-group debate will be. A valid inference from this excerpt is that this small group (1U1, 1U62, 1U29, and 1U55) engaged in dialogue in order to coordinate ideas and make a group decision. To be precise, the excerpt shows that they started to produce arguments in order to come to a unanimous decision. Also, it is interesting to note that the small-group debate offered an opportunity to the students to share their points of view and criticize the argumentation of others. For example, 1U29 said that “some physicians reacted negatively to the idea of washing their hands because of laziness”. And 1U62 communicated her/his view saying: “I think that more than laziness it was like arrogance”.

To provide a deeper understanding of the key role of step 2 in our TLS, consider the following results from the pen and paper survey: the vast majority of the respondents (67/69 in Class 1; 53/55 in Class 2) considered that the small-group debate was useful to help them to make a decision (Question 5 in [Appendix 2](#)). Some of their reasons include the following: “It was useful to broaden my own perspective. Ideas that had not been considered were brought up, and it was useful to improve the quality of the argumentation”, “listening to other opinions and counterarguments helped me to reach a conclusion”, “being exposed to different perspectives and strong arguments helped me to be more open-minded”, “listening to different points of view helped me to form a better supported opinion”, and “the exchange of ideas enriched my knowledge”. Arguably, these reasons illustrate the potential benefits of the small-group debate as a key element of our TLS. However, it is alarming to find that some respondents (32/69 in Class 1; 16/55 in Class 2) never (5/69 in Class 1; 1/55 in Class 2) or infrequently (27/69 in Class 1; 15/55 in Class 2) had the opportunity to participate in small-group debates in other university courses (Question 8 in [Appendix 2](#)).

Step 3—Students’ Engagement in Argumentative Interaction through Whole-Class Debate

The fact that most of the groups made different decisions in step 2 ([Table 3](#)) contributed to the students’ lively engagement in argumentative interaction in step 3. The following is an excerpt from the overall whole-class debate in Class 1:

I: Who wants to start?

²1U60: Us [referring to her/his small group] decided “B” and “D”. He [referring to Semmelweis] did not know how to publicize his ideas in a clear way. Also, he clashed with what was already established. He did not have enough support to counter the theory which was already established.

⁵I: Did any group make a different decision?

⁶1U4: We decided “A” and “B” because the socio-political issues of the time also had a great effect.

I: Okay. So, it is also very important for you: where and at what historical moment one does science [...] I mean, surely you can imagine that it is not the same to do science here in Colombia right now than to do science in Venezuela [...] What was Semmelweis’s hypothesis?

¹⁰1U20: The existence of a kind of rotten matter that was transmitted by medical students and physicians who did not wash their hands.

I: Everyone agrees that this is a hypothesis?

1U3: This is not.

I: What was his [referring to Semmelweis] hypothesis?

1U3: Semmelweis' hypothesis was that rotten matter caused the [puerperal] fever. The fact that the physicians spread it to women was something that supported their hypothesis; it was like evidence, I would say.

¹⁸1U56: Puerperal fever was spread by infectious matter on the hands.

I: But what is it, hypothesis or evidence?

1U56: Hypothesis.

I: But did he have one or several hypotheses? [...]

1U69: He [referring to Semmelweis] had more than one [hypothesis]. For example, at the beginning he believed that it [high mortality rate] could be due to the priest who made the tour by ringing the bell [Semmelweis hypothesized that this must have had a terrifying effect on patients with fever] [...]

²⁵1U18: We [referring to her/his small group] decided "A", "B", and "D". We think that the "A" and the "D" are interrelated. If there were no such social tension within the Austro-Hungarian Empire [option A], then it would have been easier for Semmelweis to go against the dominant ideas [option D] [...].

²⁹1U41: [options] "A", "B", and "D" could be interrelated. His [referring to Semmelweis] nationality [option A] could have affected the way other scientists perceived his publication [option B] and the validity of his research [option D] [...]

The preceding excerpt confirms that the whole-class debate served as an argumentative interaction scenario in which students were provided with opportunities to communicate their decisions as well as the evidence that they used to argue these decisions. The instructor encouraged the debate (lines 1 and 5) and used 1U4 (line 6), 1U20 (line 10), and 1U56's (line 18) points of view as a springboard to promote critical reflection. Additionally, there are two aspects in this excerpt that corroborate the controversial nature of the question, "Why did Semmelweis's views not gain acceptance among many members of the medical community in his time?". The first aspect has to do with the coexistence of different decisions: "B" and "D" (1U60, line 2), "A" and "B" (1U4, line 6), and "A", "B", and "D" (1U18, line 25). We consider that this coexistence was crucial to engage students in argumentative interaction in which they had the opportunity to cultivate their open-mindedness.

The second aspect is related to the fact that though the small groups of 1U18 (line 25) and 1U41 (line 29) made the same decision ("A", "B", and "D"), there is some discrepancy between them. The evidence which supports this is that 1U18 mentions that "the "A" and the "D" are interrelated" while 1U41's opinion is that "'A", "B", and "D" could be interrelated." This discrepancy indicates that the controversial question effectively involved students in reflexive decision-making as part of their argumentative interaction process.

To provide a deeper understanding of the key role of step 3 in our TLS, consider the following results from the survey ([Appendix 2](#)): a significant number of the participants (62/69 in Class 1; 51/55 in Class 2) who answered this commented that the whole-class debate was useful for them in making a decision (Question 6 in [Appendix 2](#)). Some of their reasons include "the ideas of each group were communicated as well as the arguments that support each answer", "to be open to listen to the opinions of other classmates helped me to have a more open view, taking into account other points of view and not just having a biased opinion within my group", "the diversity of opinions was wider than in the small group", and "it allowed me to see other types of analysis and relate them to current affairs".

These reasons suggest that the whole-class debate was relevant for the students. This is a valuable outcome; however, we cannot forget that some respondents (39/69 in Class 1; 29/55 in Class 2) never (6/69 in Class 1; 4/55 in Class 2) or infrequently (33/69 in Class 1; 25/55 in

Class 2) had the opportunity to participate in whole-class debates in other university courses (Question 9 in [Appendix 2](#)). Accordingly, we should recognize that one of the main problems we have with the promotion of argumentative interaction is that whole-class debate is virtually absent from university science education. Unfortunately, the situation becomes more complicated because the majority of participants (62/69 in Class 1; 43/55 in Class 2) never (24/69 in Class 1; 15/55 in Class 2) or infrequently (38/69 in Class 1; 28/55 in Class 2) had the opportunity to debate historical scientific controversies in other university courses (Question 7 in [Appendix 2](#)).

6 Discussion and Conclusions

The purpose of this research study was to provide evidence that the case of Semmelweis and puerperal fever—a crucial historical scientific controversy—can be used as a springboard to promote university student argumentation. To guide the reader through this section, the findings are discussed in terms of the specific ways in which they answer our two research questions. Likewise, in this section, we explain how the conclusions of the study are supported by our outcomes.

First Research Question—“Does the case of Semmelweis and puerperal fever cause controversy in a group of university students?”. An overview of the outcomes of steps 1 and 4 ([Table 2](#)) showed that the controversial question effectively caused disagreements between participants. In particular, the results of step 1 ([Table 2](#)) confirm previous research findings whereby controversial questions are a major determinant of undergraduates’ introduction to argumentative practices (Andrews [2010, 2015](#); Archila [2017](#); Muller Mirza [2015](#)). Thus, we can conclude that the historical controversy behind the case of Semmelweis and puerperal fever is a promising tool for creating student discrepancy which could then be used as a platform to promote argumentation in university science education, as attested by the results shown in [Table 2](#) (step 1). It is important to recall that the case of Semmelweis and puerperal fever has been previously used to foster informed views of nature of science (Aragón-Méndez et al. [2016, 2018](#)). Therefore, the contribution of our study is that this case has been used for the first time to explicitly promote undergraduates’ argumentation.

Having concluded that the question, “Why did Semmelweis’s views not gain acceptance among many members of the medical community in his time?”, caused controversy among participants, a second point of discussion has to do with the students’ opinion about the narrative, Semmelweis and puerperal fever (Acevedo-Díaz et al. [2016](#)). That the vast majority of participants considered that this narrative was easily understandable for them can be concluded from the results of the pen and paper survey (Question 3 in [Appendix 2](#)). It shows that this narrative could be a powerful resource for offering students the opportunity to familiarize themselves with the case of Semmelweis and puerperal fever, by getting them to feel more connected to the plausible reasons for Semmelweis’s failure. Additionally, we found that very few university students, apart from the Biology of Organisms course, have ever had heard about Semmelweis’ contributions (Question 1 in [Appendix 2](#)). These results reinforce the idea that more efforts and resources should be invested in including explicit and critical reflection of historical controversies in science classrooms (Adúriz-Bravo [2014](#); Archila [2014](#); de Hosson [2011](#); de Oliveira and Mendonça [2019](#); Garritz [2013](#); Justi and Mendonça [2016](#); Nouri and McComas [2019](#); Zemplén [2011](#)).

Thirdly, as Acevedo-Díaz et al. (2016) remind us: there is no one right answer to the controversial question, “Why did Semmelweis’s views not gain acceptance among many members of the medical community in his time?”. There could in fact be strong (rational and reasonable) arguments for each possible answer. Our results (Table 2) suggest that reading the narrative, Semmelweis and puerperal fever (Acevedo-Díaz et al. 2016), could be a particularly suitable class activity for introducing students to four plausible reasons (options A, B, C, and D) in making a decision about the controversial question. That said, a valuable result of this study is that some participants concluded that many of the members of the medical community were biased because of Semmelweis’s nationality (option A in Table 2) (Dunn 2005; Gillies 2005; Henao 1999; Stewardson and Pittet 2011; Volcy 2012). Dunn (2005), Gillies (2005), Lerner (2014), Stewardson and Pittet (2011), and Volcy (2012), among others, agree that in part Semmelweis’s failure to achieve recognition for his contribution was due to his own failure to publicize his findings. This is the same conclusion arrived at by a significant number of students (option B in Table 2).

Also, some students decided that Semmelweis’s failure could be explained by the fact that his views clashed with the dominant paradigm of his day (option D in Table 2) (Ataman et al. 2013; Gillies 2005; Henao 1999; Hernández Botero 2010; Hernández Botero and Florián Pérez 2012; Kadar 2019; Persson 2009; Volcy 2012). Nonetheless, only very few students considered physicians’ reluctance to wash their hands in calcium chloride (option C in Table 2) as a plausible reason for Semmelweis’s failure (Adriaanse et al. 2000; Cropley and Cropley 2008; Lerner 2014; Stewardson and Pittet 2011; Volcy 2012). Clearly, washing their hands in calcium chloride was regarded by Semmelweis’s fellow physicians as implying that they were dirty.

Second Research Question—“Does the TLS engage a group of university students in argumentative classroom interactions?”. Many scholars (e.g., Andrews 2010, 2015; Archila et al. 2018a, b; Pabuccu and Erduran 2017; Wieman 2017) maintain that argument and debate are virtually absent from university science education. Undergraduates enrolled in university science courses should be engaged (very frequently) in activities that involve decision-making, argumentation, and argumentative interaction (such as small-group debate and whole-class debate), coming to terms with science as a rational, social, and emotional practice, instead of as a mere accumulation of scientific conceptual knowledge (Baker 2002, 2009; Schwarz and Baker 2017). Using historical controversies to promote undergraduates’ argumentation is an alternative that has not often been explored (Archila 2015b). The outcomes from Table 3 indicate that in both classes, not all the small groups made the same decision. This situation confirms the highly controversial nature of our thought-provoking question, “Why did Semmelweis’s views not gain acceptance among many members of the medical community in his time?”. Furthermore, the results from Table 3 support the claim that the existence of a diversity of proposals (options) relating to a controversy is a vital condition for engaging students in argumentative interaction. Archila (2015b, 2017) and Baker (2002, 2009) explain that much of the reason for this is that the existence of a diversity of proposals can be a valuable catalyst for authentic and meaningful argumentative interactions.

Secondly, although in the survey a significant number of participants answered that the small-group debate was useful for them to make a decision (Question 5 in Appendix 2), this is not a common practice in university education classrooms (Question 8 in Appendix 2). This outcome reflects research in argumentation in higher education which indicates that instructors rarely adopt strategies to engage undergraduates in small-group debate (Andrews 2010; Archila 2018b; Muller Mirza 2015). On this evidence, we can conclude that the outcomes

of step 2 are in agreement with previous research findings on the intimate benefit of small-group debate as a valuable opportunity for students to exchange points of view, add information to reinforce their arguments, and reason about the controversial question under discussion (Schwarz and Baker 2017).

Thirdly, as the majority of participants commented in the survey (Question 6 in [Appendix 2](#)), the whole-class debate was useful for them to make a decision. Accordingly, in light of the previous work by Archila et al. (2018a), our findings confirm that whole-class debate is a possibility, among others, to distance university science education practice from the Confucian (and typical) science educator-controlled instruction. Likewise, this corroborates previous research that has consistently shown that student voice is increased when they are engaged in argumentative classroom interaction (Archila et al. 2018a; Muller Mirza 2015; Schwarz and Baker 2017). Naturally, as a considerable number of participants commented in the survey, the problem is that instructors rarely include whole-class debate in their university courses (Question 9 in [Appendix 2](#)) and discussion about historical scientific controversies is virtually absent from the classroom (Question 7 in [Appendix 2](#)).

To summarize, the results of the four steps (Tables 2 and 3) of our TLS in which we combined (1) argumentation, (2) decision-making, (3) historical scientific controversy, (4) argumentative interaction, (5) small-group debate, and (6) whole-class debate indicate that this sequence is a contribution to research in a new type of classroom pedagogy. In particular, the case of Semmelweis and puerperal fever—a crucial historical scientific controversy—was a powerful pedagogical resource which offered students the opportunity to experience a more authentic university science education in which they made decisions and debated about a case that had occurred more than 150 years ago. In fact, this is a case that, even today, remains controversial within the academic community (Adriaanse et al. 2000; Ataman et al. 2013; Cropley and Cropley 2008; Dunn 2005; Gillies 2005; Henao 1999; Hernández Botero 2010; Hernández Botero and Florián Pérez 2012; Kadar 2019; Lerner 2014; Persson 2009; Stewardson and Pittet 2011; Volcy 2012). This is what really makes a historical scientific controversy special in comparison with other kinds of controversies.

7 Implications

The study provides research evidence for the claim that the historical controversy related to the case of Semmelweis and puerperal fever can be used as a platform to promote undergraduate students' argumentation. Hence, some implications derived from the above-mentioned discussions may be informative. First, an interdisciplinary team of education professionals from the Faculties of Science and Education should guide and support instructors on how to treat historical controversies as a vehicle, among others, to de-emphasize the usage of examination-oriented practices as merely accumulation of scientific conceptual knowledge. Second, instructors should purposefully provide undergraduates with more opportunities to experience argumentative interaction through small-group debate and whole-class debate activities. It is also important to note that the findings reported in this article have a relevant implication for curriculum change and implementation. It is necessary not only to become keenly aware that university science courses provide undergraduates with very few opportunities to practice their argumentation skills but also to consciously realize that approaches which promote university students' argumentation require more resources and time than traditional educational programs envisage.

8 Limitations and Areas for Future Research

Despite the paramount importance for medicine of Semmelweis's introduction of anti-sepsis in the nineteenth century, there has been no study to date which has used this case to promote university students' argumentation. This is a cognitive-linguistic skill recognized by the OECD (2017) as an essential component for scientifically literate citizens in twenty-first century societies. In this study, we provide evidence for the claim that the historical controversy related to the case of Semmelweis and puerperal fever can be used as a springboard to promote undergraduates' argumentation. Nevertheless, the implementation of our TLS had several limitations that must be considered in the future to improve the efficacy of using historical scientific controversies in order to promote argumentation. First of all, the biggest limitation is that the complexity of the reasons given by participants as well as the contributions and roles of each member of the small groups were not deeply explored as this was not the purpose of the study. We were just focused on determining whether or not the historical case could cause controversy among participants. We were also interested in determining whether or not the TLS would engage them in argumentative classroom interaction. Another major limitation of the research is that there was no control group in this study. It would have been useful to contrast our results with a control group, using a traditional science educator-centered instruction, to test whether or not the TLS really engages undergraduates in argumentative classroom interactions.

We designed the TLS as a single 80-minute class session. This is a relatively short time of implementation, although this weakness applies to nearly every study interested in testing the effectiveness of a TLS (Archila 2015b; Psillos and Kariotoglou 2016). Arguably, a longer duration would be necessary to provide more robust experimental evidence. Another limitation concerns the historical scientific controversy used. A minimal number of participants apart from the Biology of Organisms course had ever heard about Semmelweis' contributions. This situation could have influenced the decisions they made. Undoubtedly, implementing the TLS with undergraduates with a higher level of knowledge about the case of Semmelweis would be necessary in order to provide more robust evidence. In addition, the sample was limited to 124 participants. This is a very small sample size that does not allow us to make generalizations from our outcomes. Hence, the results and implications of our study should be considered as exploratory, preliminary, and tentative. Furthermore, we implemented the TLS only in two university science courses. It would be interesting to implement the sequence in other courses and in other universities in order to establish additional validity.

Argumentation is an important part of science. The TLS engaged students in argumentative interaction and precipitated authentic discussion on the question, "Why did Semmelweis's views not gain acceptance among many members of the medical community in his time?". We created this TLS as a realistic, unfinished, and open alternative for instructors interested in fostering students' argumentation. Therefore, instructors may incorporate other elements and thus enrich this sequence. Future research could focus on adapting the current TLS to include other higher-order thinking skills (e.g., critical thinking), different contexts (other majors, other parts of the world), and other historical scientific controversies. By the same token, future pedagogical strategies should be based on the premise that the ultimate goal in university science education is to help university students feel that their points of view have an important place in their educational process (Archila et al. 2018a).

Acknowledgments The authors are very deeply indebted to the university students at the Biology of Organisms course for allowing them to investigate the activities in which they were engaged. Funding support from the Vice-Presidency of Research and Creation, Universidad de los Andes, Bogotá, Colombia, is gratefully acknowledged.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Appendix 1: Questionnaire

Part One

Use the text of the narrative, Semmelweis and puerperal fever (Acevedo-Díaz et al. 2016), to answer the following questions.

1. What is Semmelweis's claim?
2. What is Semmelweis's evidence?
3. What are Semmelweis's arguments?

Part Two

4. In your opinion, why did Semmelweis's views not gain acceptance among many members of the medical community in his time? (more than one option is possible).
 - a. Semmelweis's nationality
 - b. Semmelweis's failure to publicize his findings
 - c. Physicians' reluctance to wash their hands in calcium chloride
 - d. Semmelweis's views clashed with the dominant paradigm
5. Why did you make that decision?

Part Three

Having concluded the small-group debate and the whole-class debate, answer the following questions.

6. In your opinion, why did Semmelweis's views not gain acceptance among many members of the medical community in his time? (more than one option is possible).
 - a. Semmelweis's nationality
 - b. Semmelweis's failure to publicize his findings
 - c. Physicians' reluctance to wash their hands in calcium chloride
 - d. Semmelweis's views clashed with the dominant paradigm

7. Why did you make that decision?

Appendix 2: Survey

1. Apart from the Biology of Organisms course, have you ever heard about Semmelweis' contributions?
 - a. Yes
 - b. No
2. Apart from the Biology of Organisms course, have you ever received instruction in argumentation?
 - a. Yes
 - b. No
3. Was the narrative, Semmelweis and puerperal fever easily understandable for you? Explain why or why not.
4. Did you have sufficient time for reading?
 - a. Yes.
 - b. No.
5. Was the small-group debate useful for you to make a decision? Explain why or why not.
6. Was the whole-class debate useful for you to make a decision? Explain why or why not.
7. How often do you have the opportunity to debate about historical scientific controversies in other university courses?
 - a. Very frequently
 - b. Fairly frequently
 - c. Infrequently
 - d. Never
8. How often do you have the opportunity to participate in small-group debates in other university courses?
 - a. Very frequently
 - b. Fairly frequently
 - c. Infrequently
 - d. Never
9. How often do you have the opportunity to participate in whole-class debates in other university courses?
 - a. Very frequently

- b. Fairly frequently
- c. Infrequently
- d. Never

References

- Acevedo-Díaz, J. A., García-Carmona, A., & Aragón-Méndez, M. M. (2016). Un Caso de historia de la ciencia Para aprender naturaleza de la ciencia: Semmelweis y la fiebre puerperal. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 13(2), 408–422.
- Adriaanse, A. H., Pel, M., & Bleker, O. P. (2000). Semmelweis: The combat against puerperal fever. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 90(2), 153–158.
- Adúriz-Bravo, A. (2014). Revisiting school scientific argumentation from the perspective of the history and philosophy of science. In M. R. Matthews (Ed.), *International handbook of research in history, philosophy and science teaching* (pp. 1443–1472). Dordrecht: Springer.
- Andrews, R. (2010). *Argumentation in higher education: Improving practice through theory and research*. New York: Routledge.
- Andrews, R. (2015). Critical thinking and/or argumentation in higher education. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education* (pp. 49–62). New York: Palgrave Macmillan.
- Aragón-Méndez, M. M., García-Carmona, A., & Acevedo-Díaz, J. A. (2016). Aprendizaje de estudiantes de secundaria sobre la naturaleza de la ciencia mediante el Caso histórico de Semmelweis y la fiebre puerperal. *Revista Científica*, 4(27), 302–317.
- Aragón-Méndez, M. M., Acevedo-Díaz, J. A., & García-Carmona, A. (2018). Prospective biology teachers' understanding of the nature of science through an analysis of the historical case of Semmelweis and childbed fever. *Cultural Studies of Science Education*. <https://doi.org/10.1007/s11422-018-9868-y>.
- Archila, P. A. (2013). La Argumentación y sus aportes a la enseñanza bilingüe de las ciencias. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 10(3), 406–423.
- Archila, P. A. (2014). La argumentación en la formación de profesores de química: Relaciones con la comprensión de la historia de la química. *Revista Científica*, 18(1), 50–66.
- Archila, P. A. (2015a). Uso de conectores y vocabulario espontaneo en la argumentación escrita: Aportes a la alfabetización científica. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 12(3), 402–418.
- Archila, P. A. (2015b). Using history and philosophy of science to promote students' argumentation. A teaching–learning sequence based on the discovery of oxygen. *Science & Education*, 24(9), 1201–1226.
- Archila, P. A. (2015c). Evaluating evidence from a historical chemical controversy: A study in French high school. *Asia-Pacific Forum on Science Learning and Teaching*, 16(2), 1–22.
- Archila, P. A. (2017). Using drama to promote argumentation in science education: The case of “Should’ve”. *Science & Education*, 26(3–4), 345–375.
- Archila, P. A. (2018). Evaluating arguments from a play about ethics in science: A study with medical learners. *Argumentation*, 32(1), 53–76.
- Archila, P. A., Luna-Calderón, P., & Mesa-Piñeros, M. (2017). El empleo espontáneo de conectores y vocabulario relacionado con las ciencias: Implicaciones en la argumentación escrita. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 14(1), 3–23.
- Archila, P. A., Molina, J., & Truscott de Mejía, A.-M. (2018a). Using formative assessment to promote argumentation in a university bilingual science course. *International Journal of Science Education*, 40(13), 1669–1695.
- Archila, P. A., Molina, J., & Truscott de Mejía, A.-M. (2018b). Introducing undergraduates to the nature of science through the co-construction of evolutionary trees. *Evidence from a university biology course. Research in Science Education*. <https://doi.org/10.1007/s11165-018-9758-z>.
- Archila, P. A., Molina, J., & Truscott de Mejía, A.-M. (2018c). Using bilingual written argumentation to promote undergraduates' bilingual scientific literacy: Socrative® as an immersive participation tool. *International Journal of Bilingual Education and Bilingualism*. <https://doi.org/10.1080/13670050.2018.1522293>.
- Archila, P. A., Molina, J., & Truscott de Mejía, A.-M. (2019). Promoting undergraduates' awareness of the importance of thinking critically about false or inaccurate scientific information presented in news articles. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 16(13), 3106.
- Ataman, A. D., Vatanoğlu-Lutz, E. E., & Yıldırım, G. (2013). Medicine in stamps—Ignaz Semmelweis and puerperal fever. *Journal of the Turkish-German Gynecological Association*, 14(1), 35–39.

- Baker, M. J. (2002). Argumentative interactions, discursive operations and learning to model in science. In P. Bma, M. Baker, K. Stenning, & A. Tiberghien (Eds.), *The role of communication in learning to model* (pp. 303–324). Mahwah: Lawrence Erlbaum Associates.
- Baker, M. J. (2009). Argumentative interactions and the social construction of knowledge. In N. Muller Mirza & A.-N. Perret-Clermont (Eds.), *Argumentation and education: Theoretical foundations and practices* (pp. 127–144). New York: Springer.
- Best, M., & Neuhauser, D. (2004). Ignaz Semmelweis and the birth of infection control. *BMJ Quality and Safety*, 13(3), 233–234.
- Bryman, A. (2016). *Social research methods* (5th ed.). Oxford: Oxford University Press.
- Carter, K. C., & Carter, B. R. (2017). *Childbed fever: A scientific biography of Ignaz Semmelweis*. Abingdon: Routledge.
- Cropley, A., & Cropley, D. (2008). Resolving the paradoxes of creativity: An extended phase model. *Cambridge Journal of Education*, 38(3), 355–373.
- de Hosson, C. (2011). Una controversia histórica al servicio de una situación de aprendizaje: Una reconstrucción didáctica basada en diálogo sobre los dos máximos sistemas del mundo de Galileo. *Enseñanza de las Ciencias*, 29(1), 115–126.
- de Hosson, C., & Kaminski, W. (2007). Historical controversy as an educational tool: Evaluating elements of a teaching–learning sequence conducted with the text “dialogue on the ways that vision operates”. *International Journal of Science Education*, 29(5), 617–642.
- de Oliveira, J. A., & Mendonça, P. C. C. (2019). Proposta pedagógica Para o ensino explícito de argumentação: O Caso da controvérsia histórica do gás oxigênio. *Química Nova na Escola*. <https://doi.org/10.21577/0104-8899.20160150>.
- Dunn, P. M. (2005). Ignaz Semmelweis (1818–1865) of Budapest and the prevention of puerperal fever. *Archives of Disease in Childhood - Fetal and Neonatal Edition*, 90(4), 345–348.
- Encyclopædia Britannica (2018). Online version. www.britannica.com.
- Garritz, A. (2013). Teaching the philosophical interpretations of quantum mechanics and quantum chemistry through controversies. *Science & Education*, 22(7), 1787–1807.
- Gillies, D. (2005). Hempelian and Kuhnian approaches in the philosophy of medicine: The Semmelweis case. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(1), 159–181.
- Henao, G. (1999). La fiebre puerperal. La lucha de I. F. Semmelweis en la génesis de un nuevo paradigma. *IATREIA*, 12(3), 149–156.
- Hernández Botero, J. S. (2010). Sepsis, armas de fuego y microscopios: Implicaciones de la sepsis en las revoluciones de la medicina de los siglos XVI Y XIX. *Biosalud*, 9(2), 96–111.
- Hernández Botero, J. S., & Florián Pérez, M. C. (2012). The history of sepsis from ancient Egypt to the XIX century. In L. Azevedo (Ed.), *Sepsis. An ongoing and significant challenge* (pp. 3–32). London: IntechOpen.
- Justi, R., & Mendonça, P. C. C. (2016). Discussion of the controversy concerning a historical event among pre-service teachers. Contributions to their knowledge about science, their argumentative skills, and reflections about their future teaching practices. *Science & Education*, 25(7–8), 795–822.
- Kadar, N. (2019). Rediscovering Ignaz Philipp Semmelweis (1818–1865). *American Journal of Obstetrics and Gynecology*, 220(1), 26–39.
- Lerner, B. H. (2014). Searching for Semmelweis. *The Lancet*, 383(9913), 210–211.
- Loudon, I. (1986). Deaths in childbed from the eighteenth century to 1935. *Medical History*, 30(1), 1–41.
- Loudon, I. (2013). Ignaz Phillip Semmelweis’ studies of death in childbirth. *Journal of the Royal Society of Medicine*, 106(11), 461–463.
- Mavrou, K., Douglas, G., & Lewis, A. (2007). The use of Transana as a video analysis tool in researching computer–based collaborative learning in inclusive classrooms in Cyprus. *International Journal of Research & Method in Education*, 30(2), 163–178.
- Muller Mirza, N. (2015). Can we learn through disagreements? A sociocultural perspective on argumentative interactions in a pedagogical setting in higher education. *Teaching Innovations*, 28(3), 145–166.
- Nouri, N., & McComas, W. F. (2019). History of science (HOS) as a vehicle to communicate aspects of nature of science (NOS): Multiple cases of HOS instructors’ perspectives regarding NOS. *Research in Science Education*. <https://doi.org/10.1007/s11165-019-09879-9>.
- Nuland, S. B. (2003). *The doctors’ plague. Germs, childbed fever and the strange story of Ignaz Semmelweis*. New York: W. W. Norton & Co.
- Obenchain, T. G. (2016). *Genius belabored: Childbed fever and the tragic life of Ignaz Semmelweis*. Tuscaloosa: The University of Alabama Press.
- OECD. (2017). *PISA 2015 results (volume V): Collaborative problem solving*. Paris: OECD Publishing.
- Pabuccu, A., & Erduran, S. (2017). Beyond rote learning in organic chemistry: The infusion and impact of argumentation in tertiary education. *International Journal of Science Education*, 39(9), 1154–1172.

- Persson, J. (2009). Semmelweis's methodology from the modern stand-point: Intervention studies and causal ontology. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 40(3), 204–209.
- Plantin, C. (2016). *Dictionnaire de l'argumentation. Une introduction conceptuelle aux études d'argumentation*. Lyon: ENS Éditions.
- Psillos, D. (2015). Teaching and learning sequences. In R. Gunstone (Ed.), *Encyclopedia of science education* (pp. 1035–1038). Dordrecht: Springer.
- Psillos, D., & Kariotoglou, P. (2016). Theoretical issues related to designing and developing teaching-learning sequences. In D. Psillos & P. Kariotoglou (Eds.), *Iterative design of teaching-learning sequences* (pp. 11–34). Dordrecht: Springer.
- Ruiz-Primo, M. A. (2015). Cognitive labs. In R. Gunstone (Ed.), *Encyclopedia of science education* (pp. 167–171). Dordrecht: Springer.
- Schwarz, B. B., & Baker, M. J. (2017). *Dialogue, argumentation, and education: History, theory, and practice*. New York: Cambridge University Press.
- Stewardson, A., & Pittet, D. (2011). Ignác Semmelweis—Celebrating a flawed pioneer of patient safety. *The Lancet*, 378(9785), 22–23.
- Volcy, C. (2012). La investigación antigua de la fiebre puerperal: galimatías científico y objeto de reflexión. *Iatreia*, 25(2), 174–184.
- Wiemán, C. (2017). *Improving how universities teach science. Lessons from the science education initiative*. Cambridge: Harvard University Press.
- Zemplén, G. A. (2011). History of science and argumentation in science education: Joining forces? In P. V. Kakkotas, K. S. Malamitsa, & A. A. Rizaki (Eds.), *Adapting historical knowledge production to the classroom* (pp. 129–140). Rotterdam: Sense Publishers.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Pablo Antonio Archila¹ · Jorge Molina² · Anne-Marie Truscott de Mejía³

¹ Vice-Presidency of Research and Creation, Universidad de los Andes, Cra 1 N° 18A- 12, Bogotá 111711, Colombia

² Department of Biological Sciences, Universidad de los Andes, Cra 1 N° 18A- 12, Bogotá 111711, Colombia

³ School of Education, Universidad de los Andes, Cra 1 N° 18A- 12, Bogotá 111711, Colombia